DOCUMENT RESUME

ED 461 274 FL 024 713

AUTHOR Tuggy, David
TITLE Rule-Governed Allomorphy Can Be Suppletive Also.
PUB DATE 1997-00-00
NOTE 7p.; In: Work Papers of the Summer Institute of Linguistics, University of North Dakota Session. Volume 41; see FL 024 706.
PUB TYPE Information Analyses (070)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Contrastive Linguistics; Diachronic Linguistics; English; *Language Patterns; *Language Processing; Language Research; Linguistic Theory; *Morphology (Languages); Phonology; Uncommonly Taught Languages
IDENTIFIERS *Nahuatl

ABSTRACT

Suppletion is allomorphy that is produced by retrieving from the lexicon different phonological forms of the morpheme in question. A suppletive allomorph's use may be conditioned by grammatical or phonological context, or a combination. Its use therefore has dual motivation: the fact that it is governed by grammatical rule, and that its use in certain contexts is learned. This accounts for the otherwise inexplicable persistence of rule-governed allomorphy when the conditioning environment is destroyed through diachronic change, producing apparent positive exceptions to the rule. Examples include forms of the English verb "be" and forms of the Nahuatl third person singular object prefix, in which alternate forms are extremely common and occur in a very restricted set of morphophonological contexts. (MSE)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 461 274

Rule-governed Allomorphy Can Be Suppletive Also

David Tuggy

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Robert Dooley

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

FL024713

BEST COPY AVAILABLE

Rule-governed Allomorphy Can Be Suppletive Also

David Tuggy

Commonly occurring linguistic forms, including allomorphs, tend to be learned (listed in speakers' mental lexicons) even if they are formed according to the pattern of a linguistic rule. They thus have dual motivation: the motivation given by the rule, and the suppletive motivation of their having been learned. This accounts for the otherwise inexplicable persistence of rule-governed allomorphy when the conditioning environment is destroyed through diachronic change, producing apparent positive exceptions to the rule.

Suppletion is allomorphy that is produced by retrieving from the lexicon different phonological forms of the morpheme in question.[1] A conspicuous case is the verb *be* in English, which has the suppletive forms *am*, *are*, *is*, *was*, and *were*, as well as *be*, *being*, and *been*: (/æm, aɪ, ɪz, wəz, wɻ, bi, ˈbiɪŋ, bɛn ~ bɪn/). It would be quite odd to posit a phonological rule to derive, for example, /æm/ from /bi/, or vice versa, or both of them from third form; they are clearly not phonologically-related forms. Thus, it is normally assumed that forms such as *am*, *are*, *is*, *was*, *were*, and *be* are lexically stored. It is also generally assumed that a form like *being*, since it is predictable from the combination of *be* with the suffix *-ing*, is not stored. (*Been* is in some degree predictable as well, since it clearly relates to *be* and the past participial suffix *-en*, but since it is not pronounced /bin/, or perhaps /ˈbiɛn/, it is problematical. Note, for example, that *seen* cannot be pronounced */sɛn/ or */sɪn/.)

A suppletive allomorph's use may be conditioned by grammatical context (e.g., *am* occurs in present tense with first person singular subject), or by phonological context (e.g., the suppletive form *an* occurs prevocalically, and *a* preconsonantally), or by a combination of the two.

Lexical storage is generally equated with cognitive storage in (long-term) memory, the end state of a process of learning. If that identification is taken seriously, however, it is difficult to maintain the assumption that what is predictable must not be lexically stored. What governs learning of a cognitive structure is not primarily how predictable it is—in fact more-predictable structures may actually be more likely to be learned—rather, learning is responsive to factors like cognitive prominence (salience) and, especially, repetition. With often-repeated structures learning is virtually automatic and cannot easily be avoided.

Thus, it would be perfectly plausible to posit that *being*, extremely common as it is, is learned just as thoroughly as *was* and *were*, even though it is also computable from the combination of its parts. In fact, the claim that it is not learned is thoroughly counter-intuitive. However, admission that it is learned should not (and need not) mean that its composition is therefore forgotten or becomes irrelevant. The 'generativity' of lexical rules should be seen as compatible with storage (learning) of particular forms they 'produce', and, in fact, most standard models of the lexicon now allow exactly that. It is not usually invoked for regular cases such as *being*, however, but only for cases like *been* where there is some irregularity or idiosyncrasy to be recorded. *Been* would fit

---

[1] Originally suppletive forms were those which were 'supplied' to fill in for missing forms of a root in a 'defective' (i.e., incomplete) paradigm. A typically cited case is the suppletion of the missing past tense form of *go* with the past tense of *wend*. Under that definition, *am*, *is*, *are*, *was*, and *were* are suppletive with respect to *be*, but *been* and *being* are not.

Work Papers of the Summer Institute of Linguistics, University of North Dakota Session 1997 Volume 41
On line. URL: http://www.und.nodak.edu/dept/linguistics/wp/1997Tuggy.htm
Copyright © David Tuggy.

seamlessly with the proposed analysis, of course; it is suppletive in that it is learned, together with its idiosyncratic pronunciation, but that does not mean that its status as an instance of the V + *-en* pattern is abrogated.

Thus, because a form like *being* is not idiosyncratic, but regular, it is not obvious that the form has been learned. Nevertheless, so common a form is certainly learned by all English speakers, and if psychological plausibility is a goal, one's model should reflect that fact.[2]

Similar considerations favor positing learning of common phonologically-'produced' allomorphs. Where these forms are regular they can, of course, be produced by rule without existing in the theoretical lexicon. But if the lexicon is to reflect psychological (cognitive) reality, forms should be in the lexicon to the extent they are learned, regardless of how completely they can be predicted by rule. And, since learning is automatic, given sufficient repetition, common forms can be expected to be learned.

There is evidence for this position in cases where rule-governed allomorphy persists after the environment which originally conditioned it has disappeared. The 'positive exceptions' this process produces are not uncommon.

As one example, in Nahuatl the allomorphy of the third person singular object prefix *k-* ~ *ki-* has been recognized since Carochi (1645) as due to an epenthesis; i.e., the *i* is introduced in contexts where otherwise an illegal (and difficult to pronounce) consonant cluster would result. Nahuatl does not permit tautosyllabic clusters; thus, only two-consonant clusters can occur, and they only occur word-medially. For *k-* ~ *ki-*, then, *i* is epenthesized initially before a consonant, and medially between consonants in a three-consonant cluster.

This is reflected in the paradigm shown in (1). The 'underlying form' is *k-* in each case. In the third person forms, epenthesis occurs between word-initial consonants (i.e., / # k __ C), and in the 'you pl.' form it occurs after the second of a sequence of three consonants, i.e., / nk __C. This paradigm can be duplicated with any transitive CV-initial verb (that can take human subjects).

(1)	*ni-k-neki*	'I want it'	*ti-k-neki-h*	'we want it'
	ti-k-neki	'you sg. want it'	*nan-ki-neki-h*	'you pl. want it'
	ki-neki	'third sg. wants it'	*ki-neki-h*	'they want it'

The epenthesis rule is virtually exceptionless in that there are no tautosyllabic CC clusters.[3] It accounts for numerous other alternations besides these involving *k-* ~ *ki-*. To account for the 'I', 'we', and 'you sg.' forms by deletion would require a rule with hundreds of exceptions that would apply to few morphemes other than *k-* ~ *ki-*. (The evidence is laid out in Tuggy 1981.)

The alternate forms *k-* ~ *ki-* are both extremely common (at least as common as *be* and *is* in English), and they occur in a very restricted set of morpho-phonological contexts. They always occur following a subject prefix, counting the third person zero prefix, and preceding a transitive verb stem.[4] If repetition automatically causes learning, then we should expect the forms to be

---

[2] For some forms which are much less common than *being*, or common only in the speech of subgroups, it may be that only some speakers learn them. It is thus not generally useful to ask "Is structure X lexically stored in English?" but "Which (groups of) English speakers have lexically stored structure X?" Putative theoretical distinctions (e.g., between syntax and lexicon) which depend crucially on there being a strict boundary between stored and non-stored structures, are dubious.

[3] This is true for most dialects, except in certain fast-speech situations and in some borrowed words. (Many, if not most, borrowed words are adjusted to fit the canon.) In some dialects this is less true.

[4] There are a very few prefixes that can intervene between *k-* ~ *ki-* and the transitive verb stem, among them *im-* 'plural object', *on-* 'distal, thither', and *wal-* 'hither'. They are not highly relevant to the

suppletive (i.e., learned, together with the morpho-phonological contexts in which they occur). This should not mean that the cases of *ki-* would cease to be examples of the epenthesis pattern; their epenthetic analysis would be an enrichment of them, but would not explain them away or mean that they were not lexically stored.

Evidence that this is correct comes from Tetelcingo Nahuatl, where by a historical change *nan-* 'you pl. subject' changed to *na-* (or *ne-*, cf. fn. 7). If *ki-* were not lexically stored, but only produced by epenthesis, one would expect the form **na-k-neki* to surface,[5] since there would no longer be an underlying three-consonant cluster *nkn* calling for epenthesis to break it up. **na-k-neki*'s canonical structure would exactly parallel that of the first person sg. form *ni-k-neki*. What in fact happened, however, is that people continued using *ki-* in the following context: [second person plural subject] __ C. The form is *nakineki*, not **nakneki*, and similarly for all other CV-initial transitive verbs. To try to account for these forms by epenthesis would necessitate applying it where it has no phonological motivation and where all forms except *ki-* would be exceptional; i.e., *ki-* is a positive exception, a case where the rule has apparently applied even though it should not have.[6]

If the *k-* ~ *ki-* alternation was already learned in its contexts, it had double motivation: the phonological motivation of the epenthesis pattern, and the suppletive motivation of having been learned. When the phonological environment changed, removing the phonological motivation, the suppletive motivation remained. This explains why the form is still pronounced *nakineki*.[7]

It would be quite implausible to claim that the form was always and only produced by rule right up until the time the rule's conditioning environment was destroyed, and then it suddenly,

---

discussion, though they do increase the inventory of learned structures given in the Appendix. (For example, speakers have learned a *nan-k-on-* pattern as well as the *nan-k-*V pattern mentioned there.)

[5] In Tetelcingo Nahuatl final glottals are deleted, so the always-final plural suffix *-h* has disappeared from the language.

[6] To attempt to save the rule by positing a stronger boundary between the 'you pl.' prefix and *k-* than between the other subject prefixes and *k-* would be gratuitous and would involve as much extra learning (complication of the model) as would learning *ki-* for that environment.

[7] Actually, it is better to claim that when the final *n-* of the 'you pl. subject' prefix was dropped, speakers were faced with a choice. The phonological motivation would now pressure them to pronounce *ki-* as *k-*, and the learned, suppletive pattern would pressure them to retain it as *ki-*. Choices of this sort, when one group takes one option and another takes a different one, lead to dialectal differences, and it is quite conceivable that a sub-group, or the whole group, of Tetelcingo Nahuatl speakers could have begun saying **nakneki*. However, the community as a whole has settled on *nakineki*. Note that this implies that if a conditioning environment is destroyed and forms change in accordance with the rule (in this case, if the form had become **nakneki*), *this does not mean that the previous forms were not learned*, only that in choosing between possibilities, the motivation provided by any such learning did not prevail.

A functional pressure towards disambiguation is probably relevant to the Tetelcingo case. By the time the *n*-dropping was complete, many, if not most, speakers had also changed the vowel to *e*, changing the form to *nen-* or (by dropping the *n*) *ne-*. At the same time, (historically short) *i*, which had been pronounced [ɪ] (contrasting with the historically long *i* [i]), began to be often pronounced so laxly as to approximate [e]. There would thus have been only a very slight difference in many speakers' pronunciation of 'I Verb third person' and 'you pl. Verb third person' forms of CV-initial verbs (e.g., between *nikneki* 'I want it' [nɛkˈnɛke] and **nekneki* 'you pl. want it' [nɛkˈnɛke]); and the retention of *ki-* in the latter case would help disambiguate between the pairs. This is by no means a strongly predictive sort of motivation, however; the 'we' and 'you sg.' forms of all verbs have tolerated a parallel ambiguity in all speakers' speech for decades, probably well over a century (since the final *-h* disappeared).

gratuitously, became suppletive. The proposal that both mechanisms were active before the change is much more natural.

What is arguably true for such cases as *k-* ~ *ki-* is presumably true for myriads of other common alternations where the conditioning environment does not yet happen to have been destroyed so as to expose the suppletive motivation.

Appendix

The following is a more detailed description of what I would posit for the analysis of *k-* ~ *ki-* in Tetelcingo Nahuatl.

Before the final n *of* nan- *dropped*

However it is to be described theoretically,[8] the epenthesis pattern, representing the commonality of many alternations involving the presence vs. absence of *i* was well established. It was involved (however that involvement is expressed)[9] in those alternations that had been learned, including that of *k-* ~ *ki-*. It could conceivably be extended productively to novel forms. (It is hard to document cases that were clearly novel, but that is to some extent an irrelevant fact.) This pattern clearly helped relate *k-* ~ *ki-* to each other as allomorphs of the same morpheme. (Their identical meaning and alternation with each other in paradigms also established that relationship.)

Despite the redundancy, both *k-* and *ki-* were established as forms in the lexicon. They were so established in a number of different contexts. First, their occurrence with hundreds of common verb stems was learned. Examples include *kineki, nikneki, ankineki,* etc., and parallel forms with *teki* 'cut' and *māma* 'carry on the back', and dozens of other stems. These were doubtless lexical, and the fact that the *k* or *ki* preceding the stem in each case was the third person object prefix was part of what was learned. All of these forms, and particularly those with *ki-*, were instances of the epenthesis pattern and reinforced that pattern.

Arising out of those learned forms were many intermediate-level patterns, generalizations capturing their commonality. Among them were combinations of *k-* or *ki-* with particular prefixes that they commonly occur with. For the first person singular prefix *ni-*, for instance, a family of patterns existed: *ni-k-*CV for forms like *nikneki* (the most common case), *ni-k-*V for forms like *nikasi* 'I grab it', *ni-ki-*CC for cases like *nikihta* 'I see it'. These patterns included information about both the phonological shape of the surroundings and their morphological identity. In other words, since it is invariant, speakers knew both that the preceding CV was /ni/ and that in all these contexts it meant first person singular subject; they also learned not only the nature of the following segment (whether it was a consonant or a vowel), but also that it was the beginning of a transitive verb stem. Thus, in some sense the pattern we have abbreviated as *ni-k-*V would be better written as ['first person sg. subject' | *nī*] ['third person object' | *k*] [transitive verb stem | V...], where the vertical bar separates non-phonological specifications from the phonological ones. Similarly with *nan-*, *nan-ki-*C (including *nan-ki-*CV and *nan-ki-*CC subpatterns) and *nan-k-*V patterns were established in the lexicon, including the information that the phonological string *nan-* meant second person plural subject, and that what followed *ki-* or *k-* was the beginning segment of a transitive verb stem. (Thus, again, what we have abbreviated as *nan-ki-*C might more accurately

---

[8] I personally would describe it as a phonological *schema* as described by Langacker 1987 (pp. 393-394, 407-408).

[9] It is conceptually difficult to use the *generation* metaphor of generative grammar and allow any independent existence to the products of a rule. The central claim of the proposed analysis is that such independent existence must be allowed. Langacker's notion of *sanction* or legitimization fits much more easily; schemas sanction both established and novel structures.

be represented as ['second person plural subject' | *nan*]['third person object' | *k*] [transitive verb stem | C...].) Like the learned combinations with particular verb stems which they subsume, *nan-ki*-C and *ni-ki*-CC were examples of the epenthesis pattern. And as the epenthesis pattern itself could cover both established (learned) and novel, productively formed cases, so these patterns could be, and in fact not uncommonly were, used to produce the appropriate forms of verb stems where their combination with the stem was not yet learned.

In essence, people pronounced the sequence /ki/ in *nan-ki*-C for two reasons. One was because of the phonological nature of the preceding (and following) context, which together with the epenthesis pattern said "you pronounce it /ki/ (only) when not doing so will make a bad cluster". They also did it because of the non-phonological nature of the preceding context: "you pronounce it /ki/ following the 'you pl.' prefix (unless a vowel follows)". (The *nan-k*-V [and *ni-k*-V] patterns also happen to fit as examples of a well-established deletion prevocalically.)

When the final n *of* nan- *dropped*

When the final *n* of *nan-* dropped, it dropped in all the learned (established) structures that it occurred in. (This was not an automatic necessity: in fact it hung on a little longer historically in some contexts than in others.) But crucially, when it dropped in the many combinations of *nan-ki-* with particular verb stems, and thus also in *na(n)-ki*-C, examples of the *na-ki*-CV sub-pattern ceased to be subcases of the epenthesis pattern. Speakers could have chosen to allow the epenthesis pattern to dictate the result, and would then have begun to pronounce *k* instead of *ki*. They would, in effect, have been choosing to give the deciding vote to the phonological part of the preceding context. Instead they chose to give the deciding vote to the non-phonological part, to the fact that it was the second person plural subject prefix. The rule they chose to follow was not "you pronounce it /ki/ only when not doing so will make a bad cluster", but "you pronounce it /ki/ following the 'you pl.' prefix (unless a vowel follows)".

Concluding observation

The number of structures that are posited as learned under this analysis, and the amount of redundancy involved, may well be distasteful to sensibilities trained in generative grammar, with its incessant search for the highest-level generalizations possible and its eagerness to excise sub-generalizations and especially particular forms in the name of simplicity. But first, it is psychologically plausible, fitting what we know about how people learn things. They are much more tolerant of redundancy than linguists have been. And second, it explains why forms like *ki-* do not invariably revert to their putative 'underlying form' when the environment for deriving them from that form disappears.

References

Carochi, Horacio. 1645. *Arte de la lengua mexicana.* Mexico: Juan Ruiz. (Available in various modern editions).

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar*, Vol. I, *Theoretical prerequisites.* Stanford CA: Stanford University Press.

Tuggy, David. 1981b. Epenthesis of *i* in Classical and Tetelcingo Náhuatl: Evidence for multiple analyses. *Texas Linguistic Forum* 18:223-277.

David Tuggy
PO Box 8987
Catalina, AZ 85738-0987
U.S.A.

david_tuggy@sil.org